PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/67, 15/30, 5/10, 1/21, C07K 14/445, A01K 67/027, A61K 48/00

A2

(11) International Publication Number: WO 99/20774

(43) International Publication Date: 29 April 1999 (29.04.99)

(21) International Application Number: PCT/US98/22226

(22) International Filing Date: 20 October 1998 (20.10.98)

(30) Priority Data:

60/062,592	20 October 1997 (20.10.97)	US
60/085,649	15 May 1998 (15.05.98)	US

(71) Applicant: GENZYME TRANSGENICS CORPORATION [US/US]; 5 Mountain Road, Framingham, MA 01701 (US).

(72) Inventors: CHEN, Li, How; 12 Wachuset Drive, Acton, MA 01720 (US). MEADE, Harry; 62 Grasmere Street, Newton, MA 01720 (US).

(74) Agents: KEOWN, Wayne, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS

(57) Abstract

The invention provides modified recombinant nucleic acid sequences (preferably DNA) and methods for increasing the mRNA levels and protein expression of malarial surface protein MSP-1 which is known to be difficult to express in cell culture systems, mammalian cell culture systems, or in transgenic animals. The preferred protein candidates for expression using the recombinant techniques of the invention are MSP-1 proteins expressed from DNA coding sequences comprising reduced overall AT content or AT rich regions and/or mRNA instability motifs and/or rare codons relative to the native MSP-1 gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS

BACKGROUND OF THE INVENTION

Field of the invention

The invention relates to heterologous gene expression. More particularly, the invention relates to the expression of malaria genes in higher eukaryote cell systems.

Summary of the related art

Recombinant production of certain heterologous gene products is often difficult in *in vitro* cell culture systems or *in vivo* recombinant production systems. For example, many researchers have found it difficult to express proteins derived from bacteria, parasites and virus in cell culture systems different from the cell from which the protein was originally derived, and particularly in mammalian cell culture systems. One example of a therapeutically important protein which has been difficult to produce by mammalian cells is the malaria merozoite surface protein (MSP-1).

Malaria is a serious heath problem in tropical countries. Resistance to existing drugs is fast developing and a vaccine is urgently needed. Of the number of antigens that get expressed during the life cycle of *P. falciparum*, MSP-1 is the most extensively studied and promises to be the most successful candidate for vaccination. Individuals exposed to *P. falciparum* develop antibodies against MSP-1, and studies have shown that there is a correlation between a naturally acquired immune response to MSP-1 and reduced malaria morbidity. In a number of studies, immunization with purified native MSP-1 or recombinant fragments of the protein has induced at least partial protection from the parasite (Diggs et al, (1993) *Parasitol. Today* 9:300-302). Thus MSP-1 is an important target for the

development of a vaccine against *P. falciparum*.

MSP-1 is a 190-220 kDA glycoprotein. The C-terminal region has been the focus of recombinant production for use as a vaccine. However, a major problem in developing MSP-1 as a vaccine is the difficulty in obtaining recombinant proteins in bacterial or yeast expression systems that are equivalent in immunological potency to the affinity purified native protein (Chang et al., (1992) *J. Immunol.* **148**:548-555.) and in large enough quantities to make vaccine production feasible.

Improved procedures for enhancing expression of sufficient quantities of MSP-1 would be advantageous.

BRIEF SUMMARY OF THE INVENTION

The present invention provides improved recombinant DNA compositions and procedures for increasing the mRNA levels and protein expression of the malarial surface antigen MSP-1 in cell culture systems, mammalian cell culture systems, or in transgenic mammals. The preferred protein candidate for expression in an expression system in accordance with the invention is a C-terminal derivative of MSP-1 having a DNA coding sequence with reduced AT content, and eliminated mRNA instability motifs and rare codons relative to the recombinant expression systems. Thus, in a first aspect, the invention provides a DNA sequence derived from the sequence shown in SEQ ID NO 2. This derivative sequence is shown in SEQ ID NO 1.

In a second aspect, the invention provides a process for preparing a modified nucleic acid of the invention comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of the mammary gland tissue, all by replacing specific codons in the natural gene with codons recognizable to, and preferably preferred by mammary gland tissue and which code for the same amino acids as the replaced codon. This aspect of the invention further includes modified nucleic acids prepared according to the process of the invention.

In a third aspect, the invention also provides vectors comprising modified MSP-1 nucleic acids of the invention and a goat beta casein promoter and signal sequence, and host cells transformed with nucleic acids of the invention.

In a fourth aspect, the invention provides transgenic non-human mammals whose germlines comprise a nucleic acid of the invention.

In a fifth aspect, the invention provides a DNA vaccine comprising a modified MSP-1 gene according to the invention.

DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts the cDNA sequence of MSP-1_{42} modified in accordance with the invention [SEQ ID NO 1] in which 306 nucleotide positions have been replaced to lower AT content and eliminate mRNA instability motifs while maintaining the same protein amino acid sequence of MSP-1_{42}. The large letters indicate nucleotide substitutions.

Fig. 2 depicts the nucleotide sequence coding sequence of the "wild type" or native MSP- 1_{42} [SEQ ID NO 2].

Fig 3 is a codon usage table for wild type MSP-1_{42} (designated "MSP wt" in the table) and the new modified MSP-1_{42} gene (designated "edited MSP" in the table) and several milk protein genes (casein genes derived from goats and mouse). The numbers in each column indicate the actual number of times a specific codon appears in each of the listed genes. The new MSP-1_{42} synthetic gene was derived from the mammary specific codon usage by first choosing GC rich codons for a given amino acid combined with selecting the amino acids used most frequently in the milk proteins.

Fig. 4a-c depict MSP-1_{42} constructs GTC 479, GTC 564, and GTC 627, respectively as are described in the examples.

Fig. 5 panel A is a Northern analysis wherein construct GTC627 comprises the new MSP-1_{42} gene modified in accordance with the invention, GTC479 is the construct comprising the native MSP-1_{42} gene, and construct GTC469 is a negative control DNA

Fig 5 panel B is a Western analysis wherein the eluted fractions after affinity purifications numbers are collected fractions. The results show that fractions from

GTC679 the modified MSP-1_{42} synthetic gene construct reacted with polyclonal antibodies to MSP-1 and the negative control GTC479 did not.

Fig 6 depicts the nucleic acid sequences of OT1 [SEQ ID NO 3], OT2 [SEQ ID NO 4], MSP-8 [SEQ ID ON 5], MSP-2 [SEQ ID NO 6] and MSP1 [SEQ ID NO 7] described in the Examples.

Fig 7 is a schematic representation of plasmid BC574.

Fig 8 is a schematic representation of BC620.

Fig 9 is a schematic representation of BC670.

Fig 10 is a representation of a Western blot of MSP transgenic milk.

Fig 11 is a schematic representation of the nucleotide sequence of MSP42-2 [SEQ ID NO 8].

Fig 12 is a schematic representation of the BC-718.

Fig 13 is a representation of a Western blot of BC-718 expression in transgenic milk.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The patent and scientific literature referred to herein establishes the knowledge that is available to those with skill in the art. The issued US patents, allowed applications, published foreign applications, and references cited herein are hereby incorporated by reference. Any conflicts between these references and the present disclosure shall be resolved in favor of the present disclosure.

The present invention provides improved recombinant DNA compositions and procedures for increasing the mRNA levels and protein expression of the malarial surface antigen MSP-1 in cell culture systems, mammalian cell culture systems, or in transgenic mammals. The preferred protein candidate for expression in an expression system in accordance with the invention is a C-terminal derivative of MSP-1 having a DNA coding sequence with reduced AT content, and eliminated mRNA instability motifs and rare codons relative to the recombinant expression systems. Thus, in a first aspect, the invention provides a DNA sequence derived from the sequence shown in SEQ ID NO 2. This derivative sequence is shown in SEQ ID NO 1.

In preferred embodiments, the nucleic acid of the invention is capable of expressing MSP-1 in mammalian cell culture systems, or in transgenic mammals at a level which is at least 25%, and preferably 50% and even more preferably at least 100% or more of that expressed by the natural gene in mammalian cell culture systems, or in transgenic mammals under identical conditions.

As used herein, the term "expression" is meant mRNA transcription resulting in protein expression. Expression may be measured by a number of techniques known in the art including using an antibody specific for the protein of interest. By "natural gene" or "native gene" is meant the gene sequence, or fragments thereof (including naturally occurring allelic variations), which encode the wild type form of MSP-1 and from which the modified nucleic acid is derived. A "preferred codon "means a codon which is used more prevalently by the cell or

tissue type in which the modified MSP-1 gene is to be expressed, for example, in mammary tissue. Not all codon changes described herein are changes to a preferred codon, so long as the codon replacement is a codon which is at least recognized by the mouse mammary tissue. The term "reduced AT content" as used herein means having a lower overall percentage of nucleotides having A (adenine) or T (thymine) bases relative to the natural MSP-1 gene due to replacement of the A or T containing nucleotide positions or A and/or T containing codons with nucleotides or codons recognized by mouse mammary tissue and which do not change the amino acid sequence of the target protein.

In a second aspect, the invention provides a process for preparing a modified nucleic acid of the invention comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of mammary gland tissue, all by replacing specific codons in the natural gene with codons recognizable to, and prferably preferred by mammary gland tissue and which code for the same amino acids as the replaced codon. Standard reference works describing the general principals of recombinant DNA technology include Watson, J.D. et al, *Molecular Biology of the Gene*, Volumes I and II the Benjamin/Cummings Publishing Company, Inc. publisher, Menlo Park, CA (1987) Darnell, J.E. et al., *Molecular Cell Biology*, Scientific American Books, Inc., Publisher, New York, NY (1986); Old, R.W., et al., *Principles of Gene Manipulation: An Introduction to Genetic Engineering*, 2d edition, University of California Press, publisher, Berkeley CA (1981); Maniatis, T., et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory, publisher, Cold Spring Harbor, NY (1989) and *Current Protocols in Molecular Biology*, Ausubel et al., Wiley Press, New York, NY (1992). This aspect of the invention further includes modified nucleic acids prepared according to the process of the invention.

Without being limited to any theory, previous research has indicated that a

conserved AU sequence (AUUUA) from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation (Shaw, G. and Kamen, R. *Cell* 46:659-667). The focus in the past has been on the presence of these instability motifs in the untranslated region of a gene. The instant invention is the first to recognize an advantage to eliminating the instability sequences in the coding region of the MSP-1 gene.

In a third aspect, the invention also provides vectors comprising modified MSP-1 nucleic acids of the invention and a goat beta casein promoter and signal sequence, and host cells transformed with nucleic acids of the invention.

In a fourth aspect, the invention provides transgenic non-human mammals whose germlines comprise a nucleic acid of the invention. General principals for producing transgenic animals are known in the art. See for example Hogan et al., *Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory, (1986); Simons et al, *Bio/Technology* 6:179-183, (1988); Wall et al., *Biol. Reprod.* 32:645-651, (1985); Buhler et al., *Bio/Technology*, 8:140-143 (1990); Ebert et al., *Bio/Technology* 9:835-838 (1991); Krimenfort et al., *Bio/Technology* 9:844-847 (1991); Wall et al., *J.Cell. Biochem.* 49:113-120 (1992). Techniques for introducing foreign DNA sequences into mammals and their germ cells were originally developed in the mouse. See e.g., Gordon et al., *Proc. Natl. Acad. Sci. USA* 77:7380-7384, (1980); Gordon and Ruddle, *Science* 214: 1244-1246 (1981); Palmiter and Brinster, *Cell* 41: 343-345, 1985; Brinster et al., *Proc Natl. Acad Sci., USA* 82:4438-4442 (1985) and Hogan et al. (*ibid.*). These techniques were subsequently adapted for use with larger animals including cows and goats. Up until very recently, the most widely used procedure for the generation of transgenic mice or livestock, several hundred linear molecules of the DNA of interest in the form of a transgenic expression construct are injected into one of the pro- nuclei of a fertilized egg. Injection of DNA into the cytoplasm of a zygote is also widely used. Most recently cloning of an entire

transgenic cell line capable of injection into an unfertilized egg has been achieved (KHS Campbell et al., *Nature* 380 64-66, (1996)).

The mammary gland expression system has the advantages of high expression levels, low cost, correct processing and accessibility. Known proteins, such as bovine and human alpha- lactalbumin have been produced in lactating transgenic animals by several researchers. (Wright et al, *Bio/Technology* 9:830-834 (1991); Vilotte et al, *Eur. J. Biochem.*,186:43-48 (1989); Hochi et al., *Mol Reprod. And Devel.* 33:160-164 (1992); Soulier et al., *FEBS Letters* 297(1,2):13-18 (1992)) and the system has been shown to produce high levels of protein.

In a fifth aspect, the invention provides a DNA vaccine comprising a modified MSP-1 gene according to the invention. Such DNA vaccines may be delivered without encapsulation, or they may be delivered as part of a liposome, or as part of a viral genome. Generally, such vaccines are delivered in an amount sufficient to allow expression of the modified MSP-1 gene and to elicit an antibody response in an animal, including a human, which receives the DNA vaccine. Subsequent deliveries, at least one week after the first delivery, may be used to enhance the antibody response. Preferred delivery routes include introduction via mucosal membranes, as well as parenteral administration.

Examples

Creation of novel modified MSP-1_{42} gene

A novel modified nucleic acid encoding the C-terminal fragment of MSP-1 is provided. The novel, modified nucleic acid of the invention encoding a 42 kD C-terminal part of MSP-1 (MSP-1_{42}) capable of expression in mammalian cells of the invention is shown in Fig. 1. The natural MSP-1_{42} gene (Fig 2) was not capable of being expressed in mammalian cell culture or in transgenic mice Analysis of the natural MSP- 1_{42} gene suggested several characteristics that distinguish it from

mammalian genes. First, it has a very high overall AT content of 76%. Second, the mRNA instability motif, AUUUA, occurred 10 times in this 1100 bp DNA segment (Fig 2). To address these differences a new MSP-1_{42} gene was designed. Silent nucleotide substitution was introduced into the native MSP-1_{42} gene at 306 positions to reduce the overall AT content to 49.7%. Each of the 10 AUUUA mRNA instability motifs in the natural gene were eliminated by changes in codon usage as well. To change the codon usage, a mammary tissue specific codon usage table, Fig. 3a, was created by using several mouse and goat mammary specific proteins. The table was used to guide the choice of codon usage for the modified MSP-1_{42} gene as described above. For example as shown in the Table in Fig. 3a, in the natural gene, 65% (25/38) of the Leu was encoded by TTA, a rare codon in the mammary gland. In the modified MSP-1_{42} gene, 100% of the Leu was encoded by CTG, a preferred codon for Leu in the mammary gland.

An expression vector was created using the modified MSP-1_{42} gene by fusing the first 26 amino acids of goat beta-casein to the N-terminal of the modified MSP-1_{42} gene and a SalI-Xho I fragment which carries the fusion gene was subcloned into the XhoI site of the expression vector pCDNA3. A His6 tag was fused to the 3' end of the MSP-1_{42} gene to allow the gene product to be affinity purified. This resulted in plasmid GTC627 (Fig.4c).

To compare the natural MSP-1_{42} gene construct to the modified MSP-1_{42} nucleic acid of the invention, an expression vector was also created for the natural MSP-1_{42} gene and the gene was added to mammalian cell culture and injected into mice to form transgenic mice as follows:

Construction of the native MSP-1_{42} *Expression Vector*

To secrete the truncated merozoite surface protein-1 (MSP-1) of Plasmodium falciparum, the wild type gene encoding the 42KD C-terminal part of MSP-1 ($MSP\text{-}1_{42}$) was fused to either the DNA sequence that encodes the first 15 or the first 26 amino acids of the goat beta-casein. This is achieved by first PCR amplify the MSP-1 plasmid (received from Dr. David Kaslow, NIH) with primers MSP1 and MSP2 (Fig. 6), then cloned the PCR product into the TA vector (Invitrogen). The BglII-XhoI fragments of the PCR product was ligated with oligos OT1 and OT2 (Fig. 6) into the expression vector pCDNA3. This yielded plasmid GTC564 (Fig.4b), which encodes the 15 amino acid beta- casein signal peptide and the first 11 amino acids of the mature goat beta-casein followed by the native $MSP\text{-}1_{42}$ gene. Oligos MSP-8 and MSP-2 (Fig. 6) were used to amplify MSP-1 plasmid by PCR, the product was then cloned into TA vector. The XhoI fragment was exercised and cloned into the XhoI site of the expression vector pCDNA3 to yield plasmid GTC479 (Fig.4a), which encoded 15 amino acid goat beta-casein signal peptide fused to the wild-type $MSP\text{-}1_{42}$ gene. A His6 tag was added to the 3' end of $MSP\text{-}1_{42}$ gene in GTC 564 and GTC 479.

Native $MSP\text{-}1_{42}$ Gene Is Not Expressed In COS-7 Cells

Expression of the native MSP gene in cultured COS-7 cells was assayed by transient transfection assays. GTC479 and GTC564 plasmids DNA were introduced into COS-7 cells by lipofectamine (Gibco-BRL) according to manufacturer's protocols. Total cellular RNA was isolated from the COS cells two days post-transfection. The newly synthesized proteins were metabolically labeled for 10 hours by adding ^{35}S methionine added to the culture media two days-post transfection.

To determine the MSP mRNA expression in the COS cells, a Northern blot was probed with a ^{32}P labeled DNA fragment from GTC479. No MSP RNA was detected in GTC479 or GTC564 transfectants (data not shown). Prolonged exposure revealed residual levels of degraded MSP mRNA. The ^{35}S labeled culture supernatants and

the lysates were immunoprecipitated with a polyclonal antibody raised against MSP. Immunoprecipitation experiments showed that no expression from either the lysates or the supernatants of the GTC479 or GTC564 transfected cells (data not shown). These results showed that the native MSP-1 gene was not expressed in COS cells.

Native MSP-1_{42} Gene is Not Expressed in the Mammary Gland of Transgenic Mice

The SalI-XhoI fragment of GTC479, which encoded the 15 amino acids of goat beta-casein signal peptide, the first 11 amino acids of goat beta-casein, and the native MSP-1_{42} gene, was cloned into the XhoI site of the beta-casein expressed in vector BC350. This yielded plasmid BC574 (Fig.7). A SalI-NotI fragment of BC574 was injected into the mouse embryo to generate transgenic mice. Fifteen lines of transgenic mice were established. Milk from the female founder mice was collected and subjected to Western analysis with polycolonal antibodies against MSP. None of the seven mice analyzed were found to express MSP-1_{42} protein in their milk. To further determine if the mRNA of MSP-1_{42} was expressed in the mammary gland, total RNA was extracted from day 11 lactating transgenic mice and analyzed by Northern blotting. No MSP-1_{42} mRNA was detected by any of the BC 574 lines analyzed. Therefore, the MSP-1_{42} transgene was not expressed in the mammary gland of transgenic mice. Taken together, these experiments suggest that native parasitic MSP-1_{42} gene could not be expressed in mammalian cells, and the block is as the level of mRNA abundance.

Expression of MSP in the Mammalian Cells

Transient transfection experiments were performed to evaluate the expression of

the modified MSP-1_{42} gene of the invention in COS cells. GTC627 and GTC479 DNA were introduced into the COS-7 cells. Total RNA was isolated 48 hours post-transfection for Northern analysis. The immobilized RNA was probed with ^{32}P labeled SalI-XhoI fragment of GTC627. A dramatic difference was observed between GTC479 and GTC627. While no MSP-1_{42} mRNA was detected in the GTC479 transfected cells as shown previously, abundant MSP-1_{42} mRNA was expressed by GTC627 (Fig. 5, Panel A). GTC 469 was used as a negative control and comprises the insert of GTC564 cloned into cloning vector PU19, a commercially available cloning vector. A metabolic labeling experiment with ^{35}S methionine followed by immunoprecipitation with polyclonal antibody (provided by D. Kaslow NIAID, NIH) against MSP showed that MSP-1_{42} protein was synthesized by the transfected COS cells (Fig.5, Panel B). Furthermore, MSP-1_{42} was detected in the transfected COS supernatant, indicating the MSP-1_{42} protein was also secreted. Additionally, using Ni-NTA column, MSP-1_{42} was affinity purified from the GTC627 transfected COS supernatant.

These results demonstrated that the modification of the parasitic MSP-1_{42} gene lead to the expression of MSP mRNA in the COS cells. Consequently, the MSP-1_{42} product was synthesized and secreted by mammalian cells.

Polyclonal antibodies used in this experiment may also be prepared by means well known in the art (*Antibodies: A Laboratory Manual*, Ed Harlow and David Lane, eds. Cold Spring Harbor Laboratory, publishers (1988)). Production of MSP serum antibodies is also described in Chang et al., *Infection and Immunity* (1996) 64:253-261 and Chang et al., (1992) *Proc Natl. Acad. Sci. USA* 86:6343-6347.

The results of this analysis indicate that the modified MSP-1_{42} nucleic acid of the invention is expressed at a very high level compared to that of the natural protein

which was not expressed at all. These results represent the first experimental evidence that reducing the AT % in a gene leads to expression of the MSP gene in heterologous systems and also the first evidence that removal of AUUUA mRNA instability motifs from the MSP coding region leads to the expression of MSP protein in COS cells. The results shown in Fig. 5, Panel A Northern (i.e. no RNA with native gene and reasonable levels with a modified DNA sequence in accordance with the invention), likely explains the increase in protein production.

The following examples describe the expression of MSP1-42 as a native non-fusion (and non-glycosylated) protein in the milk of transgenic mice.

Construction of MSP Transgene

To fuse MSP1-42 to the 15 amino acid ß-casein signal peptide, a pair of oligos, MSP203 and MSP204 (MSP203: ggccgctcgacgccaccatgaaggtcctcataattgcc tgtctggtggctctggccattgcagccgtcactccctccgtcat, MSP204: cgatgacggagggagtgacggctg caatggccagagccaccagacaggcaattatgaggaccttcatggtggcgtcgagc), which encode the 15 amino acid - casein signal and the first 5 amino acid of the MSP1-42 ending at the Cla I site, was ligated with a Cla I-Xho I fragment of BC620 (Fig. 8) which encodes the rest of the MSP1-42 gene, into the Xho I site of the expression vector pCDNA3. A Xho I fragment of this plasmid (GTC669) was then cloned into the Xho I site of milk specific expression vector BC350 to generate B670 (Fig.9)

Expression of MSP1-42 in the milk of transgenic mice

A Sal I-Not I fragment was prepared from plasmid BC670 and microinjected into the mouse embryo to generate transgenic mice. Transgenic mice was identified by extracting mouse DNA from tail biopsy followed by PCR analysis using oligos GTC17 and MSP101 (sequences of oligos: GTC17, GATTGACAAGTAATACGCTGTTTCCTC, Oligo MSP101, GGATTCAATAGATACGG). Milk from the female founder transgenic mice was collected at day 7 and day 9 of lactation, and subjected to western analysis to determine the expression level of MSP-1-42 using an polyclonal anti-MSP antibody and monoclonal anti MSP antibody 5.2 (Dr. David Kaslow, NIH). Results indicated that the level of MSP-1-42 expression in the milk of

transgenic mice was at 1-2 mg/ml (Fig. 10).

Construction of MSP1-42 glycosylation sites minus mutants

Our analysis of the milk produced MSP revealed that the transgenic MSP protein was N-glycosylated. To eliminate the N-glycosylation sites in the MSP1-42 gene, Asn. (N) at positions 181 and 262 were substituted with Gln.(Q). The substitutions were introduced by designing DNA oligos that anneal to the corresponding region of MSP1 and carry the AAC to CAG mutations. These oligos were then used as PCR primers to produce DNA fragments that encode the N to Q substitutions.

To introduce N262-Q mutation, a pair of oligos, MSPGYLYCO-3 (CAGGGAATGCTGCAGATCAGC) AND MSP42-2 (AATTCTCGAGTTAGTG GTGGTGGTGGTGGTGATCGCAGAAAATACCATG, FIG. 11), were used to PCR amplify plasmid GTC627, which contains the synthetic MSP1-42 gene. The PCR product was cloned into pCR2.1 vector (Invitrogen). This generated plasmid GTC716.

To introduce N181-Q mutation, oligos MSPGLYCO-1 (CTCCTTGTTCAGG AACTTGTAGGG) and MSPGLCO-2 (GTCCTGCAGTACACATATGAG, Fig 4) were used to amplify plasmid GTC 627. The PCR product was cloned into pCR2.1. This generated plasmid GTC700.

The MSP double glycosylation mutant was constructed by the following three steps: first, a Xho I-Bsm I fragment of BC670 and the Bsm 1-Xho I fragment of GTC716 is ligated into the Xho I site of vector pCR2.1. This resulted a plasmid that contain the MSP-1-42 gene with N262-Q mutation. EcoN I-Nde I fragment of this plasmid was then replaced by the EcoN I-Nde I fragment from plasmid GTC716 to introduce the second mutation, N181-Q. A Xho I fragment of this plasmid was finally cloned into BC350 to generate BC718 (Fig. 12).

Transgenic expression of nonglycosylated MSP-1

BC718 has the following characteristics: it carries the MSP1-42 gene under the control of the ß-casein promoter so it can be expressed in the mammary gland of the transgenic animal during lactation. Further, it encodes a 15 amino acid ß-casein leader sequence fused directly to MSP1-42.

so that the MSP1-42, without any additional amino acid at its N-terminal, can be secreted into the milk. Finally, because the N-Q substitutions, the MSP produced in the milk of the transgenic animal by this construct will not be N-glycosylated. Taken together, the transgenic MSP produced in the milk by BC718 is the same as the parasitic MSP.

A SalI/XhoI fragment was prepared from plasmid BC718 and microinjected into mouse embryos to generate transgenic mice. Transgenic animals were identified as described previously. Milk from female founders was collected and analyzed by Western blotting with antibody 5.2. The results, shown in Figure 13, indicate expression of nonglycosylated MSP1 at a concentration of 0.5 to 1 mg/ml.

What is claimed is:

1. A modified MSP-1 nucleic acid sequence, as shown in SEQ ID NO 1.

2. A process for preparing a modified MSP-1 nucleic acid comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of the mouse mammary gland.

3. A modified MSP-1 nucleic acid prepared by the process according to claim 2.

4. A vector comprising a modified MSP-1 nucleic acid according to claim 1 and a goat beta casein promoter and signal sequence.

5. A host cell transformed with the vector according to claim 4.

6. A transgenic non-human mammal comprising a vector according to claim 4.

7. A transgenic non-human mammal comprising a modified MSP-1 nucleic acid according to claim 1.

8. A DNA vaccine comprising a modified nucleic acid according to claim 1.

FIG. 1

```
                  C
 1 GCCGTCACTCCCCTCCCGTCA GATAACATCCTGTCCAAGATCGAGAACGAGTA
 1▶Al a Val Thr Pro Ser Val I l e Asp Asn I l e Leu Ser Lys I l e Gl u Asn Gl u Tyr G
56 AGGTGCTGTACCCTGAAGCCCTGGCAGGGGTCTACCGGAGCCTGAAGAAGCAG
19▶ l uVal Leu Tyr Leu Lys Pro Leu Al a Gl y Val Tyr Arg Ser Leu Lys Lys Gl n
109 CTGGAGAACAACGTGATGACCTTCAACGTGAACGTGAAGGATATCCTGAACAGC
37▶ Leu Gl u Asn Asn Val Met Thr Phe Asn Val Asn Val Lys Asp I l e Leu Asn Ser
163 CGGTTCAACAAGCGGGAGAACTTCAAGAACGTGCTGGAGCAGCGATCTGATCCC
55▶ Arg Phe Asn Lys Arg Gl u Asn Phe Lys Asn Val Leu Gl u Ser Asp Leu I l e Pr
216 CTACAAGGATCTGACCAGCAGCAACTACGTGGTCAAGGATCCCTACAAGTTCC
72▶ oTyr Lys Asp Leu Thr Ser Ser Asn Tyr Val Val Lys Asp Pro Tyr Lys Phe L
269 TGAACAAGGAGAAGAGAGATAAGTTCCTGAGCAGTTACAACTACATCAAGGATAG
90▶ euAsn Lys Gl u Lys Arg Asp Lys Phe Leu Ser Ser Tyr Asn Tyr I l e Lys Asp Se
324 CATTGATACCGATATCAACTTCGCCAACGATGTCCTGGATACTACAAGATCCT
108▶ r I l e Asp Thr Asp I l e Asn Phe Al a Asn Asp Val Leu Gl y Tyr Tyr Lys I l e Le
378 GTCCGAGAAGTACAAGAGCGATCTGGATTCAATCAAGAAGTACATCAACGATAA
126▶ u Ser Gl u Lys Tyr Lys Ser Asp Leu Asp Ser I l e Lys Lys Tyr I l e Asn Asp Ly
432 GCAGGGAGAGAACGAGAAGTACCTGCCCTTCCTGAACAACATCGAGACCCTGTA
144▶ s Gl n Gl y Gl u Asn Gl u Lys Tyr Leu Pro Phe Leu Asn Asn I l e Gl u Thr Leu Ty
486 CAAGACCGTCAACGATAAGATTGATCTGTTCGTGATCCACCTGGAGGCCAAGGT
162▶ r Lys Thr Val Asn Asp Lys I l e Asp Leu Phe Val I l e Hi s Leu Gl u Al a Lys Va
                  NdeI
540 CCTGAACTACACATATGAGAACGAGCAACGTGGAGGTCAAGATCAAGGAGCTGAA
180▶ I Leu Asn Tyr Thr Tyr Gl u Lys Ser Asn Val Gl u Val Lys I l e Lys Gl u Leu As
594 TTACCTGAAGACCATCCAGGATAAGCTGGCCGATTTCAAGAAGAACAACAACTT
198▶ nTyr Leu Lys Thr I l e Gl n Asp Lys Leu Al a Asp Phe Lys Lys Asn Asn Asn Ph
648 CGTCGGGATCGCCGATCTGAGCACCGATTACAACCACAACAACCTGCTGACCAA
216▶ eVal Gl y I l e Al a Asp Leu Ser Thr Asp Tyr Asn Hi s Asn Asn Leu Leu Thr Ly
702 GTTCCTGAGCACCGGTATGGTCTTCGAAAACCTGGCCAAGACCGTCCTGAGCAA
234▶ sPhe Leu Ser Thr Gl y Met Val Phe Gl u Asn Leu Al a Lys Thr Val Leu Ser As
756 CCTGCTGGATGGGAACCTGCAGGGATGCTGAACATCAGCCAGCACCAGTGTGT
252▶ n Leu Leu Asp Gl y Asn Leu Gl n Gl y Met Leu Asn I l e Ser Gl n Hi s Gl n Cys Va
810 GAAGAAGCAGTGTCCCCAGAACAGCGGTGTTTCAGACACCTGGATGAGAGAGA
270▶ I Lys Lys Gl n Cys Pro Gl n Asn Ser Gl y Cys Phe Arg Hi s Leu Asp Gl u Arg Gl
864 GGAGTGTAAGTGTCTGCTGAACTACAAGCAGGAAGGTGATAAGTGTGTGGAAAAC
288▶ u Gl u Cys Lys Cys Leu Leu Asn Tyr Lys Gl n Gl u Gl y Asp Lys Cys Val Gl u Asn
919 CCCAATCCTACTTGTAACGAGAACAATGGTGGATGTGATGCCGATGCCAAGTGTACCG
307▶ Pro Asn Pro Thr Cys Asn Gl u Asn Asn Gl y Gl y Cys Asp Al a Asp Al a Lys Cys Thr G
977 AGGAGGATTCAGGGAGCAACGGGAAGAAGATCACCTGTGAGTGTACCAAGCCTGATT
326▶ I u Gl u Asp Ser Gl y Ser Asn Gl y Lys Lys I l e Thr Cys Gl u Cys Thr Lys Pro Asp S
1034 CTTATCCACTGTTCGATGGTATCTTCTGTAGT
345▶ er Tyr Pro Leu Phe Asp Gl y I l e Phe Cys Ser
```

F16.2


1 GCAGTAACTCCTTCCGTAATT .AACATACTTTCTAAAATTGAAAATGAATA
1 AlaValThrProSerVal I I e ?AsnI I eLeuSerLysI I eGluAsnGluTy
EcoNI (73)
56 AGGTTTTATATTTAAAACCTTTAGCAGGTGTTTATAGAAGTTTAAAAAAACAATT
19 luValLeuTyrLeuLysProLeuAlaGlyValTyrArgSerLeuLysLysGlnLe
111 AGAAAATAACGTTATGACATTTAATGTTAATGTTAAGGATATTTTAAATTCACGA
37 uGluAsnAsnValMetThrPheAsnValAsnValLysAspI leLeuAsnSerArg
166 TTTAATAAACGTGAAAATTTCAAAAATGTTTTAGAATCAGATTTAATTCCATATA
56 PheAsnLysArgGluAsnPheLysAsnValLeuGluSerAspLeuI leProTyrL
221 AAGATTTAACATCAAGTAATTATGTTGTCAAAGATCCATATAAATTTCTTAATAA
74 ysAspLeuThr Ser SerAsnTyrVal Val LysAspProTyrLysPheLeuAsnLy
276 AGAAAAAAGAGATAAATTCTTAAGCAGTTATAATTATATTAAGGATTCAATAGAT
92 sGluLysArgAspLysPheLeuSer Ser TyrAsnTyrI leLysAspSer I leAsp
331 ACGGATATAAATTTTGCAAATGATGTTCTTGGATATTATAAAATATTATCCGAAA
111 ThrAspI leAsnPheAlaAsnAspVal LeuGlyTyrTyrLysI leLeuSer GluL
386 AATATAAATCAGATTTAGATTCAATTAAAAAATATATCAACGACAAACAAGGTGA
129 ysTyrLysSerAspLeuAspSer I leLysLysTyrI leAsnAspLysGlnGlyGl
441 AAATGAGAAATACCTTCCCTTTTTAAACAATATTGAGACCTTATATAAAACAGTT
147 uAsnGluLysTyrLeuProPheLeuAsnAsnI leGluThr LeuTyrLysThr Val
496 AATGATAAAATTGATTTATTTGTAATTCATTTAGAAGCAAAAGTTCTAAATTATA
166 AsnAspLysI leAspLeuPheVal I leHisLeuGluAlaLysVal LeuAsnTyrT
551 CATATGAGAAATCAAACGTAGAAGTTAAAATAAAAGAACTTAATTACTTAAAAAC
184 hrTyrGluLysSerAsnVal GluVal LysI leLysGluLeuAsnTyrLeuLysTh
606 AATTCAAGACAAATTGGCAGATTTTAAAAAAAATAACAATTTCGTTGGAATTGCT
202 r I leGlnAspLysLeuAlaAspPheLysLysAsnAsnAsnPheVal GlyI leAla
661 GATTTATCAACAGATTATAACCATAATAACTTATTGACAAAGTTCCTTAGTACAG
221 AspLeuSer ThrAspTyrAsnHisAsnAsnLeuLeuThr LysPheLeuSer Thr G
716 GTATGGTTTTTGAAAATCTTGCTAAAACCGTTTTATCTAATTTACTTGATGGAAA
239 lyMetVal PheGluAsnLeuAlaLysThr Val LeuSerAsnLeuLeuAspGlyAs
771 CTTGCAAGGTATGTTAAACATTTCACAACACCAATGCGTAAAAAAACAATGTCCA
257 nLeuGlnGlyMetLeuAsnI leSer GlnHisGlnCysVal LysLysGlnCysPro
826 CAAAATTCTGGATGTTTCAGACATTTAGATGAAAGAGAAGAATGTAAATGTTTAT
276 GlnAsnSer GlyCysPheArgHisLeuAspGluArgGluGluCysLysCysLeuL
881 TAAATTACAAACAAGAAGGTGATAAATGTGTTGAAAATCCAAATCCTACTTGTAA
294 euAsnTyrLysGlnGluGlyAspLysCysVal GluAsnProAsnProThr CysAs
936 CGAAAATAATGGTGGATGTGATGCAGATGCCAAATGTACCGAAGAAGATTCAGGT
312 nGluAsnAsnGlyGlyCysAspAlaAspAlaLysCysThr GluGluAspSer Gly
991 AGCAACGGAAAGAAAATCACATGTGAATGTACTAAACCTGATTCTTATCCACTTT
331 SerAsnGlyLysLysI leThr CysGluCysThr LysProAspSer TyrProLeuP
PstI (1059)
1046 TCGATGGTATTTTCTGCAGTCACCACCACCACCACCACTAACT
349 heAspGlyI lePheCysSer HisHisHisHisHisHis***

Fig. 3A

Codon	AA	goat b-casein	goat K-casei	SP wt	Edited MSP	mouse b-casei	mouse a-casein	mouse	sein	mouse e-casei
TTT	Phe	5	4	8	0	4	8	3		4
TTC	Phe	4	3	7	15	4	6	7		1
TTA	Leu	0	2	25	0	0	0	0		0
TTG	Leu	0	2	3	0	0	0	0		1
TCT	Ser	5	1	4	1	13	5	7		5
TCC	Ser	2	2	2	3	6	14	8		2
TCA	Ser	1	4	10	1	1	3	2		0
TCG	Ser	0	1	0	0	0	0	0		0
TAT	Tyr	2	7	17	2	1	3	2		1
TAC	Tyr	1	2	3	18	2	6	6		7
TAA	***	1	2	0	0	1	0	1		0
TAG	***	0	0	0	0	0	0	0		0
TGT	Cys	1	1	10	12	0	0	1		0
TGC	Cys	0	2	2	0	2	2	2		1
TGA	***	0	0	0	0	0	1	0		1
TGG	Trp	1	1	0	0	0	2	2		2
CTT	Leu	9	1	9	0	16	9	3		3
CTC	Leu	5	2	0	0	7	8	0		1
CTA	Leu	1	2	1	0	1	2	1		0
CTG	Leu	11	5	0	38	10	17	4		1
CCT	Pro	17	6	4	2	8	6	3		0
CCC	Pro	12	0	1	6	8	6	6		4
CCA	Pro	3	13	5	1	5	6	2		2
CCG	Pro	1	1	0	1	0	0	0		1
CAT	His	0	1	3	0	2	6	2		1
CAC	His	5	3	1	4	4	0	3		0
CAA	Gln	5	9	9	0	9	21	9		7
CAG	Gln	16	6	0	9	21	32	12		8
CGT	Arg	0	1	1	0	0	0	0		0
CGC	Arg	0	0	0	0	1	0	0		0
CGA	Arg	0	0	1	0	0	0	0		1
CGG	Arg	1	0	0	3	0	0	0		0
ATT	Ile	4	5	13	0	3	4	3		4
ATC	Ile	6	3	2	20	7	5	8		5
ATA	Ile	1	3	5	0	1	0	2		0
ATG	Met	7	3	3	3	4	12	2		13
ACT	Thr	7	6	3	2	6	5	1		4
ACC	Thr	2	7	3	13	4	4	4		4
ACA	Thr	2	4	9	1	1	1	2		0
ACG	Thr	0	0	1	0	0	0	2		0
AAT	Asn	2	6	29	3	4	6	3		1
AAC	Asn	2	3	12	38	4	9	4		6
AAA	Lys	7	6	38	0	6	7	3		5
AAG	Lys	6	4	4	42	3	6	13		7
AGT	Ser	2	6	5	2	3	6	6		5
AGC	Ser	5	0	2	16	2	6	6		3
AGA	Arg	2	2	4	3	1	8	1		1
AGG	Arg	0	2	0	0	0	0	0		1
GTT	Val	5	6	15	0	7	4	2		3
GTC	Val	8	2	1	11	7	3	3		0
GTA	Val	2	2	5	0	2	4	1		3
GTG	Val	8	4	0	10	6	3	5		3
GCT	Ala	1	3	2	0	8	17	4		2
GCC	Ala	4	7	1	8	6	3	3		3
GCA	Ala	3	7	6	1	4	13	1		1
GCG	Ala	0	1	0	0	0	0	0		0
GAT	Asp	4	5	25	27	3	6	4		2
GAC	Asp	0	2	2	0	1	2	1		3
GAA	Glu	10	6	21	3	6	12	9		6
GAG	Glu	9	5	4	22	5	5	5		5
GGT	Gly	2	1	8	4	0	0	0		0
GGC	Gly	0	0	0	0	0	0	0		0
GGA	Gly	2	1	6	3	1	0	1		0
GGG	Gly	1	0	0	7	1	0	0		0

Page 1

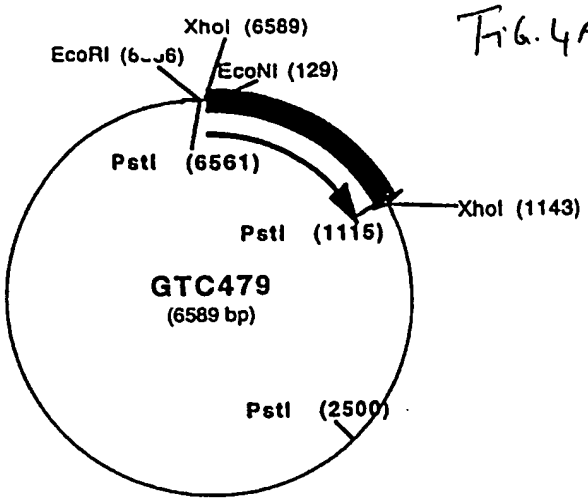


Fig. 4A

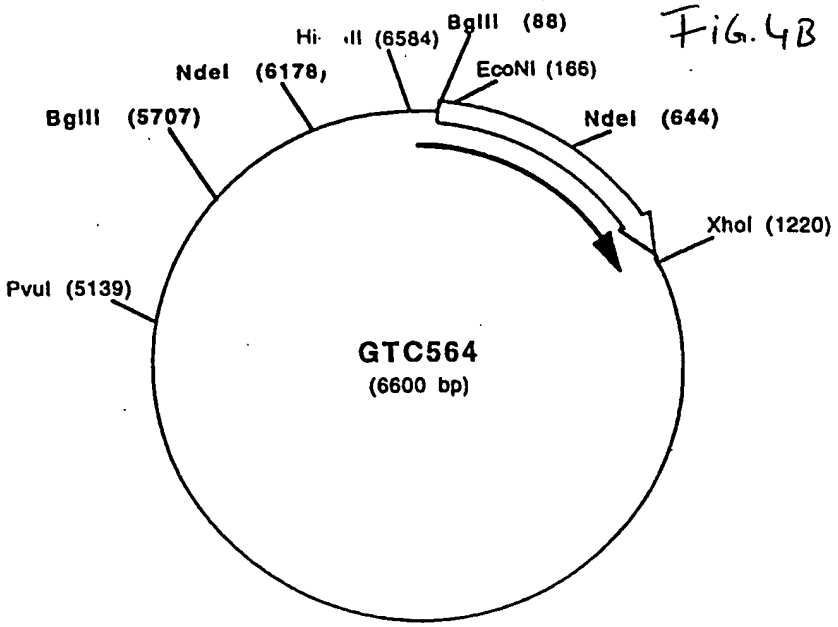

Hi· ·II (6584)
BglII (88)
NdeI (6178)
EcoNI (166)
BglII (5707)
NdeI (644)
XhoI (1220)
PvuI (5139)
GTC564
(6600 bp)

Fig. 4B

FIG. 4C

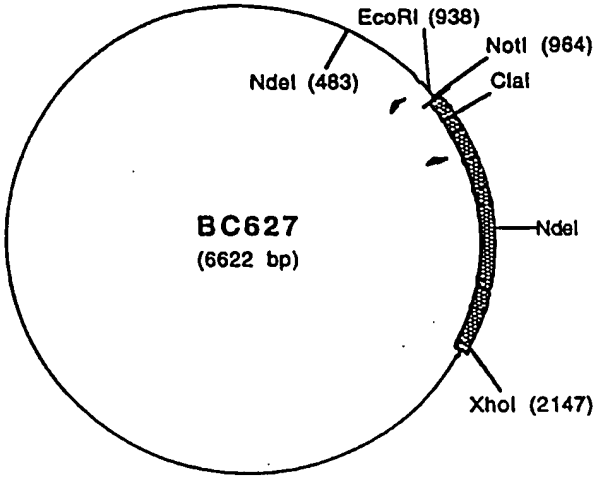


FIG. 5

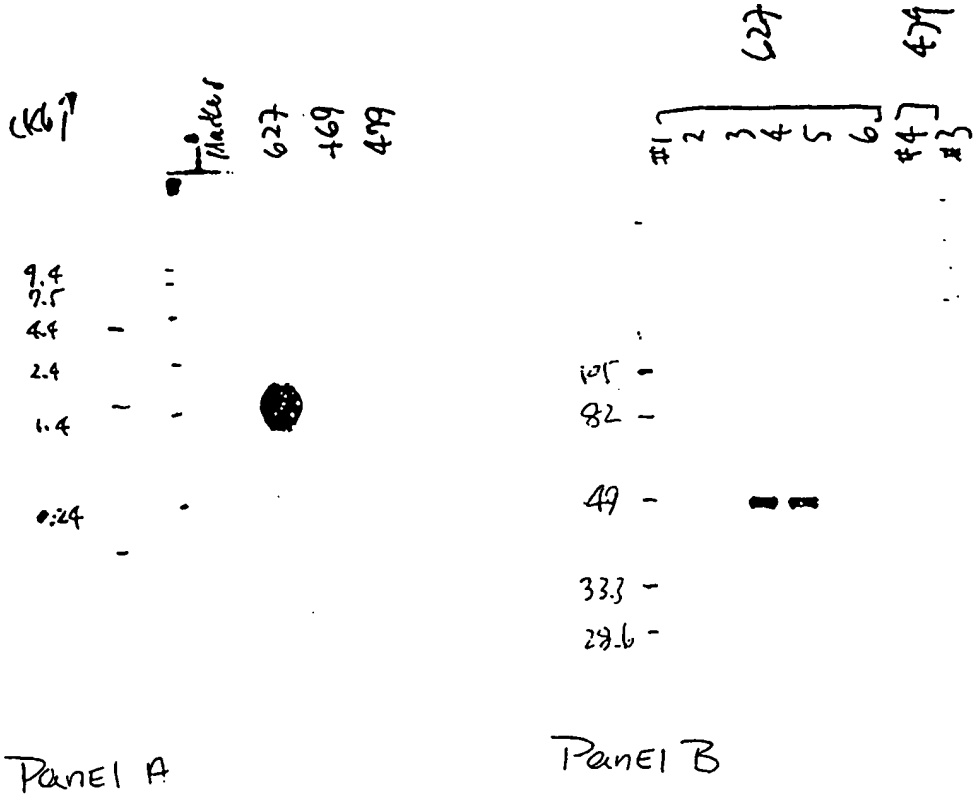


FIG. 6

Oligos used:

OT1:
TCG ACG AGA GCC ATG AAG GTC CTC ATC CTT GCC TGT CTG GTG GCT CTG GCC ATT GCA AGA GAG CAG GAA GAA CTC AAT GTA GTC GGT A,

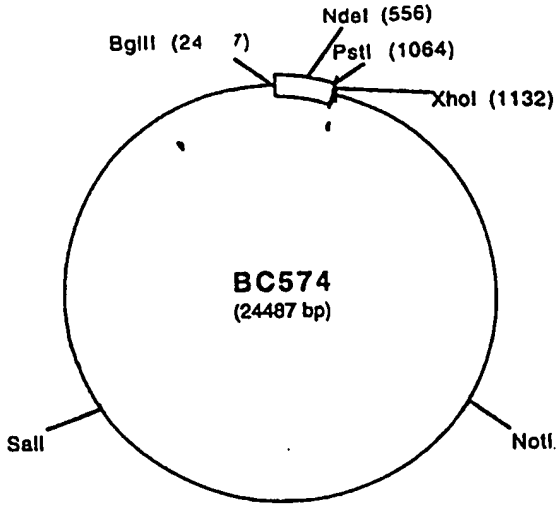
OT2:
GAT CTA CCG ACT ACA TTG AGT TCT TCC TGC TCT CTT GCA ATG GCC AGA GCC ACC AGA CAG GCA AGG ATG AGG ACC TTC ATG GCT CTC G,

MSP1:
AATAGATCTGCAGTAACTCCTTCCGTAATTG,

MSP2:
AATTCTCGAGTTAGTGGTGGTGGTGGTGGTGACTGCAGAAATACCATC

MSP8:
TAACTCGAGCGAACCATGAAGGTCCTCATCCTTGCCTGTCTGGTGGCTCTGG CCATTGCA

FIG. 7


NdeI (556)
BglII (24 7)
PstI (1064)
XhoI (1132)
BC574
(24487 bp)
SalI
NotI

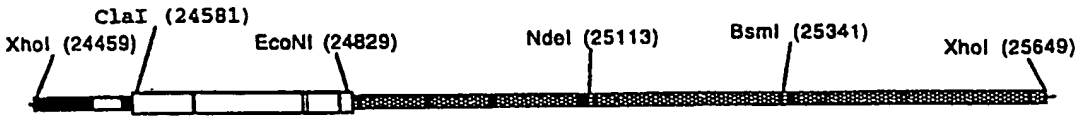


Diagram of BC620

FIGURE 8

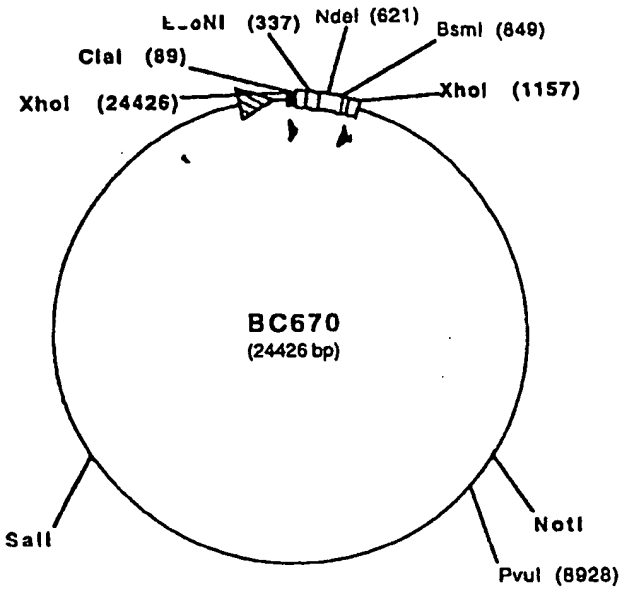
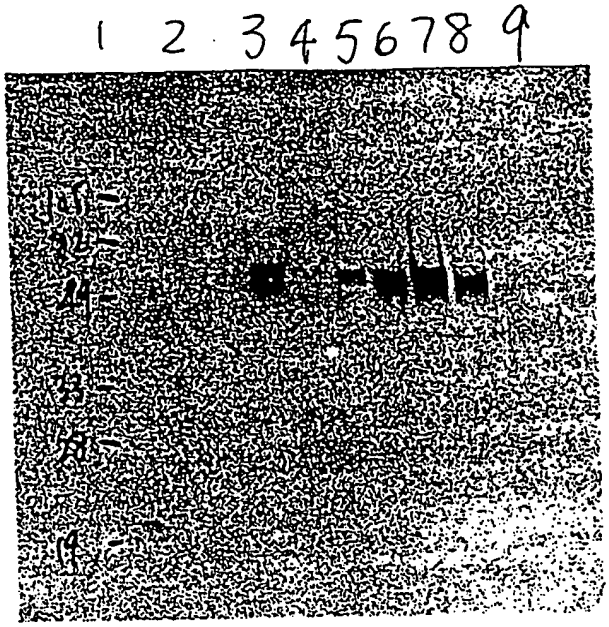

EcoNI (337)
NdeI (621)
BsmI (849)
ClaI (89)
XhoI (24426)
XhoI (1157)
BC670
(24426 bp)
SalI
NotI
PvuI (8928)

FIGURE 9



Western Analysis of MSP transgenic milk.
Lane 1, Molecular weight marker; lane 2, nontransgenic mice milk; lane 3, milk from BC628-146 transgenic mouse; lane 4-9, milk from BC670 transgenic mice. The blot was reacted with monoclonal antibody 5.2 against MSP.

FIGURE 10

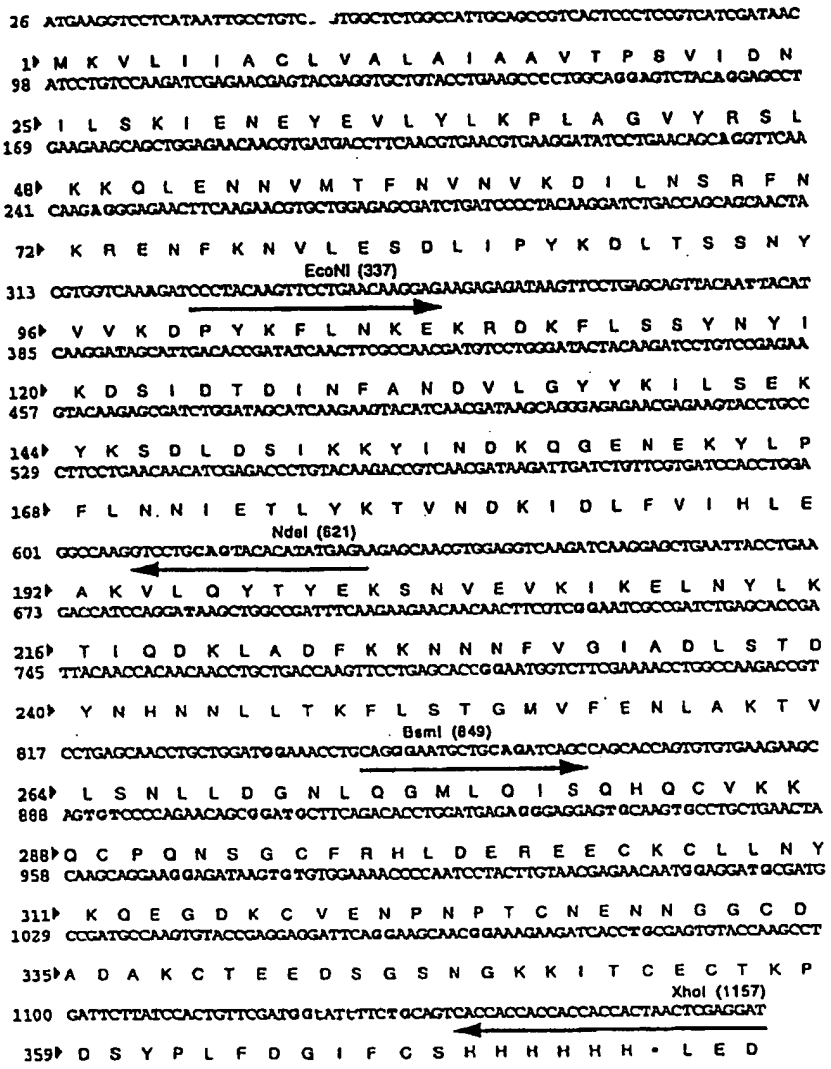


FIGURE 11

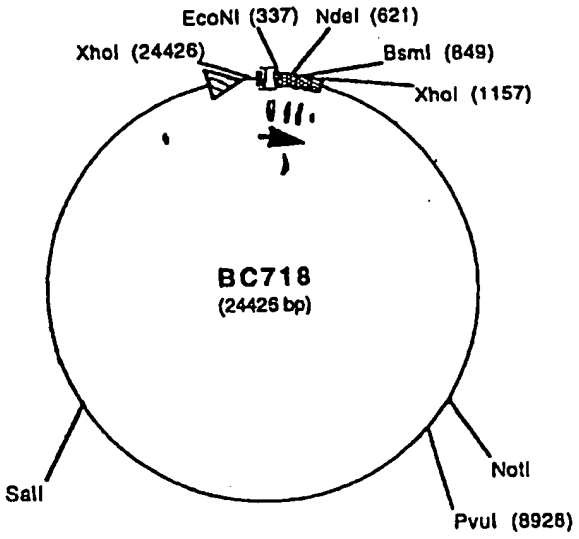

EcoNI (337)
NdeI (621)
XhoI (24426)
BsmI (849)
XhoI (1157)
BC718
(24426 bp)
SalI
NotI
PvuI (8928)

FIGURE 12

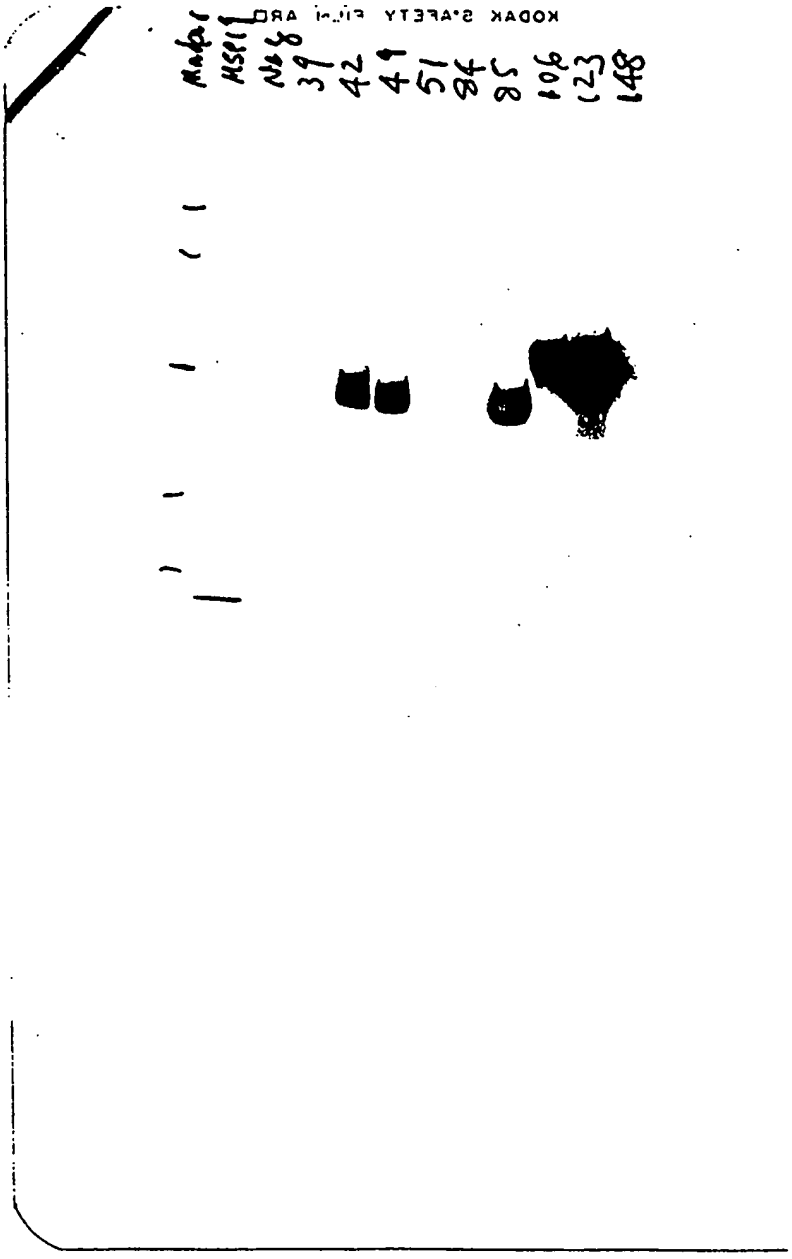


FIGURE 13

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/67, 15/30, 5/10, 1/21, C07K 14/445, A01K 67/027, A61K 48/00

A3

(11) International Publication Number: WO 99/20774

(43) International Publication Date: 29 April 1999 (29.04.99)

(21) International Application Number: PCT/US98/22226

(22) International Filing Date: 20 October 1998 (20.10.98)

(30) Priority Data:

60/062,592	20 October 1997 (20.10.97)	US
60/085,649	15 May 1998 (15.05.98)	US

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

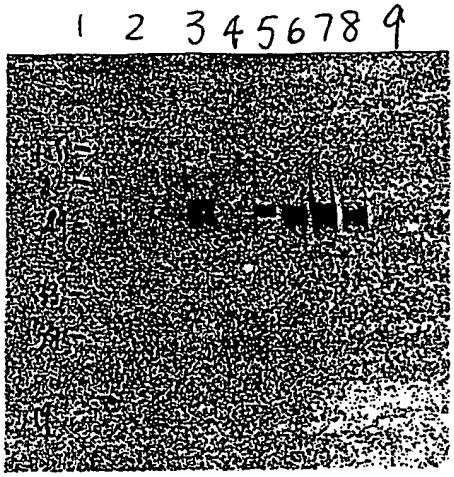
(88) Date of publication of the international search report:
26 August 1999 (26.08.99)

(71) Applicant: GENZYME TRANSGENICS CORPORATION [US/US]; 5 Mountain Road, Framingham, MA 01701 (US).

(72) Inventors: CHEN, Li, How; 12 Wachuset Drive, Acton, MA 01720 (US). MEADE, Harry; 62 Grasmere Street, Newton, MA 01720 (US).

(74) Agents: KEOWN, Wayne, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).

(54) Title: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS



Western Analysis of MSP transgenic milk.
Lane 1, Molecular weight marker; lane 2, nontransgenic mice milk; lane 3, milk from BC628-146 transgenic mouse; lane 4-9, milk from BC670 transgenic mice. The blot was reacted with monoclonal antibody 5.2 against MSP.

(57) Abstract

The invention provides modified recombinant nucleic acid sequences (preferably DNA) and methods for increasing the mRNA levels and protein expression of malarial surface protein MSP-1 which is known to be difficult to express in cell culture systems, mammalian cell culture systems, or in transgenic animals. The preferred protein candidates for expression using the recombinant techniques of the invention are MSP-1 proteins expressed from DNA coding sequences comprising reduced overall AT content or AT rich regions and/or mRNA instability motifs and/or rare codons relative to the native MSP-1 gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Int. tional Application No
PCT/US 98/22226

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/67 C12N15/30 C12N5/10 C12N1/21 C07K14/445
A01K67/027 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A01K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 98 14583 A (PAN WEIQING ;BUJARD HERMANN (DE); TOLLE RALF (DE)) 9 April 1998 (1998-04-09) abstract examples 1,6 claims	1-8
A	WO 94 28930 A (VIROGENETICS CORP) -22 December 1994 (1994-12-22) abstract	1-8
A	EP 0 682 115 A (MYCOGEN PLANT SCIENCE INC) 15 November 1995 (1995-11-15) the whole document	

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 July 1999	15/07/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lejeune, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Ir · ational Application No
PCT/US 98/22226

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 31115 A (DAVIES MARY ELLEN ;PERRY HELEN C (US); SHIVER JOHN W (US); FREED D) 28 August 1997 (1997-08-28) abstract page 50 - page 59	
A	WO 91 08216 A (GENPHARM INT) 13 June 1991 (1991-06-13) abstract	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

In. ational Application No
PCT/US 98/22226

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9814583 A	09-04-1998	DE 19640817 A	14-05-1998
		AU 4864997 A	24-04-1998
WO 9428930 A	22-12-1994	US 5766597 A	16-06-1998
		AU 701825 B	04-02-1999
		AU 7060294 A	03-01-1995
		CA 2168124 A	22-12-1994
		EP 0717636 A	26-06-1998
EP 0682115 A	15-11-1995	AT 132193 T	15-01-1996
		AU 623429 B	14-05-1992
		AU 4118289 A	15-03-1990
		CN 1044298 A	01-08-1990
		DE 68925253 D	08-02-1996
		EP 0359472 A	21-03-1990
		ES 2083384 T	16-04-1996
		GR 3019280 T	30-06-1996
		JP 2186989 A	23-07-1990
		US 5380831 A	10-01-1995
		US 5567600 A	22-10-1996
		US 5567862 A	22-10-1996
WO 9731115 A	28-08-1997	AU 2124697 A	10-09-1997
		CZ 9802667 A	17-03-1999
		EP 0904380 A	31-03-1999
		HR 970092 A	30-04-1998
		NO 983876 A	21-10-1998
		PL 328730 A	15-02-1999
WO 9108216 A	13-06-1991	AT 140027 T	15-07-1996
		AU 656720 B	16-02-1995
		AU 6960891 A	26-06-1991
		CA 2075206 A	02-06-1991
		CN 1053446 A	31-07-1991
		DE 69027683 D	08-08-1996
		DE 69027683 T	13-02-1997
		DK 502976 T	11-11-1996
		EP 0502976 A	16-09-1992
		EP 0737746 A	16-10-1996
		ES 2090299 T	16-10-1996
		FI 923485 A	31-07-1992
		GR 3021239 T	31-01-1997
		OA 9669 A	15-05-1993
		RU 2095414 C	10-11-1997
		US 5633076 A	27-05-1997
		US 5741957 A	21-04-1998